

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANTS: van Rooijen, Gijs
Alcantara, Joenel
Moloney, Maurice M.
- (ii) TITLE OF INVENTION: Method for Cleavage of Fusion Proteins
- (iii) NUMBER OF SEQUENCES: 4
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: BERESKIN & PARR
(B) STREET: 40 King Street West
(C) CITY: Toronto
(D) STATE: Ontario
(E) COUNTRY: Canada
(F) ZIP: M5H 3Y2
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 09/402,488
(B) FILING DATE: 23-APR-1998
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Gravelle, Micheline
(B) REGISTRATION NUMBER: 40,261
(C) REFERENCE/DOCKET NUMBER: 9369-98
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (416) 364-7311
(B) TELEFAX: (416) 361-1398

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1096 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
- (ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..1029

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG TCC CCT ATA CTA GGT TAT TGG AAA ATT AAG GGC CTT GTG CAA CCC

09 JUN 12 AM 7:21

RECEIVED
GENE 1660/2900

Met	Ser	Pro	Ile	Leu	Gly	Tyr	Trp	Lys	Ile	Lys	Gly	Leu	Val	Gln	Pro	
1				5					10					15		
ACT	CGA	CTT	CTT	TTG	GAA	TAT	CTT	GAA	GAA	AAA	TAT	GAA	GAG	CAT	TTG	96
Thr	Arg	Leu	Leu	Leu	Glu	Tyr	Leu	Glu	Glu	Lys	Tyr	Glu	Glu	His	Leu	
			20					25					30			
TAT	GAG	CGC	GAT	GAA	GGT	GAT	AAA	TGG	CGA	AAC	AAA	AAG	TTT	GAA	TTG	144
Tyr	Glu	Arg	Asp	Glu	Gly	Asp	Lys	Trp	Arg	Asn	Lys	Lys	Phe	Glu	Leu	
		35					40					45				
GGT	TTG	GAG	TTT	CCC	AAT	CTT	CCT	TAT	TAT	ATT	GAT	GGT	GAT	GTT	AAA	192
Gly	Leu	Glu	Phe	Pro	Asn	Leu	Pro	Tyr	Tyr	Ile	Asp	Gly	Asp	Val	Lys	
	50					55					60					
TTA	ACA	CAG	TCT	ATG	GCC	ATC	ATA	CGT	TAT	ATA	GCT	GAC	AAG	CAC	AAC	240
Leu	Thr	Gln	Ser	Met	Ala	Ile	Ile	Arg	Tyr	Ile	Ala	Asp	Lys	His	Asn	
65					70					75					80	
ATG	TTG	GGT	GGT	TGT	CCA	AAA	GAG	CGT	GCA	GAG	ATT	TCA	ATG	CTT	GAA	288
Met	Leu	Gly	Gly	Cys	Pro	Lys	Glu	Arg	Ala	Glu	Ile	Ser	Met	Leu	Glu	
				85					90					95		
GGA	GCG	GTT	TTG	GAT	ATT	AGA	TAC	GGT	GTT	TCG	AGA	ATT	GCA	TAT	AGT	336
Gly	Ala	Val	Leu	Asp	Ile	Arg	Tyr	Gly	Val	Ser	Arg	Ile	Ala	Tyr	Ser	
			100					105					110			
AAA	GAC	TTT	GAA	ACT	CTC	AAA	GTT	GAT	TTT	CTT	AGC	AAG	CTA	CCT	GAA	384
Lys	Asp	Phe	Glu	Thr	Leu	Lys	Val	Asp	Phe	Leu	Ser	Lys	Leu	Pro	Glu	
		115					120					125				
ATG	CTG	AAA	ATG	TTC	GAA	GAT	CGT	TTA	TGT	CAT	AAA	ACA	TAT	TTA	AAT	432
Met	Leu	Lys	Met	Phe	Glu	Asp	Arg	Leu	Cys	His	Lys	Thr	Tyr	Leu	Asn	
	130					135					140					
GGT	GAT	CAT	GTA	ACC	CAT	CCT	GAC	TTC	ATG	TTG	TAT	GAC	GCT	CTT	GAT	480
Gly	Asp	His	Val	Thr	His	Pro	Asp	Phe	Met	Leu	Tyr	Asp	Ala	Leu	Asp	
145					150					155				160		
GTT	GTT	TTA	TAC	ATG	GAC	CCA	ATG	TGC	CTG	GAT	GCG	TTC	CCA	AAA	TTA	528
Val	Val	Leu	Tyr	Met	Asp	Pro	Met	Cys	Leu	Asp	Ala	Phe	Pro	Lys	Leu	
				165					170					175		
GTT	TGT	TTT	AAA	AAA	CGT	ATT	GAA	GCT	ATC	CCA	CAA	ATT	GAT	AAG	TAC	576
Val	Cys	Phe	Lys	Lys	Arg	Ile	Glu	Ala	Ile	Pro	Gln	Ile	Asp	Lys	Tyr	
			180					185					190			
TTG	AAA	TCC	AGC	AAG	TAT	ATA	GCA	TGG	CCT	TTG	CAG	GGC	TGG	CAA	GCC	624
Leu	Lys	Ser	Ser	Lys	Tyr	Ile	Ala	Trp	Pro	Leu	Gln	Gly	Trp	Gln	Ala	
		195					200					205				
ACG	TTT	GGT	GGT	GGC	GAC	CAT	CCT	CCA	AAA	TCG	GAT	CTG	GTT	CCG	CGT	672
Thr	Phe	Gly	Gly	Gly	Asp	His	Pro	Pro	Lys	Ser	Asp	Leu	Val	Pro	Arg	
	210					215					220					
GGA	TCC	CCG	AAT	TCC	CGG	GTC	GAC	TCG	AGC	GGC	CGC	GCT	GAG	ATC	ACC	720
Gly	Ser	Pro	Asn	Ser	Arg	Val	Asp	Ser	Ser	Gly	Arg	Ala	Glu	Ile	Thr	
225					230					235				240		
AGG	ATC	CCT	CTG	TAC	AAA	GGC	AAG	TCT	CTG	AGG	AAG	GCG	CTG	AAG	GAG	768

Arg	Ile	Pro	Leu	Tyr	Lys	Gly	Lys	Ser	Leu	Arg	Lys	Ala	Leu	Lys	Glu		
				245					250					255			
CAT	GGG	CTT	CTG	GAG	GAC	TTC	CTG	CAG	AAA	CAG	CAG	TAT	GGC	ATC	AGC		816
His	Gly	Leu	Leu	Glu	Asp	Phe	Leu	Gln	Lys	Gln	Gln	Tyr	Gly	Ile	Ser		
			260					265					270				
AGC	AAG	TAC	TCC	GGC	TTC	GTC	GTC	TAT	ACC	GAC	TGT	ACC	GAG	TCC	GGT		864
Ser	Lys	Tyr	Ser	Gly	Phe	Val	Val	Tyr	Thr	Asp	Cys	Thr	Glu	Ser	Gly		
		275					280					285					
CAG	AAC	CTC	TGT	CTC	TGT	GAG	GGT	TCC	AAC	GTC	TGT	GGT	CAG	GGT	AAC		912
Gln	Asn	Leu	Cys	Leu	Cys	Glu	Gly	Ser	Asn	Val	Cys	Gly	Gln	Gly	Asn		
	290					295					300						
AAG	TGT	ATC	CTC	GGT	TCC	GAC	GGT	GAG	AAG	AAC	CAG	TGT	GTC	ACC	GGT		960
Lys	Cys	Ile	Leu	Gly	Ser	Asp	Gly	Glu	Lys	Asn	Gln	Cys	Val	Thr	Gly		
305					310				315					320			
GAG	GGA	ACC	CCA	AAG	CCA	CAG	TCC	CAC	AAC	GAC	GGT	GAC	TTT	GAG	GAG		1008
Glu	Gly	Thr	Pro	Lys	Pro	Gln	Ser	His	Asn	Asp	Gly	Asp	Phe	Glu	Glu		
				325				330					335				
ATC	CCA	GAG	GAG	TAT	CTC	CAG	TAAAGATCTA	AGCTTGCTGC	TGCTATCGAA								1059
Ile	Pro	Glu	Glu	Tyr	Leu	Gln											
			340														
TTCCTGCAGC	CCGGGGGATC	CACTAGTTCT	AGAGCGG														1096

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 343 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Ser	Pro	Ile	Leu	Gly	Tyr	Trp	Lys	Ile	Lys	Gly	Leu	Val	Gln	Pro		
1				5				10						15			
Thr	Arg	Leu	Leu	Leu	Glu	Tyr	Leu	Glu	Glu	Lys	Tyr	Glu	Glu	His	Leu		
		20						25					30				
Tyr	Glu	Arg	Asp	Glu	Gly	Asp	Lys	Trp	Arg	Asn	Lys	Lys	Phe	Glu	Leu		
		35					40					45					
Gly	Leu	Glu	Phe	Pro	Asn	Leu	Pro	Tyr	Tyr	Ile	Asp	Gly	Asp	Val	Lys		
	50					55					60						
Leu	Thr	Gln	Ser	Met	Ala	Ile	Ile	Arg	Tyr	Ile	Ala	Asp	Lys	His	Asn		
65					70				75					80			
Met	Leu	Gly	Gly	Cys	Pro	Lys	Glu	Arg	Ala	Glu	Ile	Ser	Met	Leu	Glu		
				85				90						95			
Gly	Ala	Val	Leu	Asp	Ile	Arg	Tyr	Gly	Val	Ser	Arg	Ile	Ala	Tyr	Ser		

100					105					110					
Lys	Asp	Phe	Glu	Thr	Leu	Lys	Val	Asp	Phe	Leu	Ser	Lys	Leu	Pro	Glu
		115					120					125			
Met	Leu	Lys	Met	Phe	Glu	Asp	Arg	Leu	Cys	His	Lys	Thr	Tyr	Leu	Asn
	130					135					140				
Gly	Asp	His	Val	Thr	His	Pro	Asp	Phe	Met	Leu	Tyr	Asp	Ala	Leu	Asp
145					150					155					160
Val	Val	Leu	Tyr	Met	Asp	Pro	Met	Cys	Leu	Asp	Ala	Phe	Pro	Lys	Leu
				165					170					175	
Val	Cys	Phe	Lys	Lys	Arg	Ile	Glu	Ala	Ile	Pro	Gln	Ile	Asp	Lys	Tyr
			180					185					190		
Leu	Lys	Ser	Ser	Lys	Tyr	Ile	Ala	Trp	Pro	Leu	Gln	Gly	Trp	Gln	Ala
		195					200					205			
Thr	Phe	Gly	Gly	Gly	Asp	His	Pro	Pro	Lys	Ser	Asp	Leu	Val	Pro	Arg
	210					215					220				
Gly	Ser	Pro	Asn	Ser	Arg	Val	Asp	Ser	Ser	Gly	Arg	Ala	Glu	Ile	Thr
225					230					235					240
Arg	Ile	Pro	Leu	Tyr	Lys	Gly	Lys	Ser	Leu	Arg	Lys	Ala	Leu	Lys	Glu
				245					250					255	
His	Gly	Leu	Leu	Glu	Asp	Phe	Leu	Gln	Lys	Gln	Gln	Tyr	Gly	Ile	Ser
			260					265					270		
Ser	Lys	Tyr	Ser	Gly	Phe	Val	Val	Tyr	Thr	Asp	Cys	Thr	Glu	Ser	Gly
		275					280					285			
Gln	Asn	Leu	Cys	Leu	Cys	Glu	Gly	Ser	Asn	Val	Cys	Gly	Gln	Gly	Asn
	290					295					300				
Lys	Cys	Ile	Leu	Gly	Ser	Asp	Gly	Glu	Lys	Asn	Gln	Cys	Val	Thr	Gly
305					310					315					320
Glu	Gly	Thr	Pro	Lys	Pro	Gln	Ser	His	Asn	Asp	Gly	Asp	Phe	Glu	Glu
				325					330					335	
Ile	Pro	Glu	Glu	Tyr	Leu	Gln									
			340												

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 819 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(ix) FEATURE:

- (A) NAME/KEY: CDS

(B) LOCATION: 1..816

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATG CGG GGT TCT CAT CAT CAT CAT CAT CAT GGT ATG GCT AGC ATG ACT	48
Met Arg Gly Ser His His His His His His Gly Met Ala Ser Met Thr	
1 5 10 15	
GGT GGA CAG CAA ATG GGT CGG GAT CTG TAC GAC GAT GAC GAT AAG GAT	96
Gly Gly Gln Gln Met Gly Arg Asp Leu Tyr Asp Asp Asp Asp Lys Asp	
20 25 30	
CCG AGC TCG AGA TCT GCA GAA ATC GGA TCC GCT GAG ATC ACC AGG ATC	144
Pro Ser Ser Arg Ser Ala Glu Ile Gly Ser Ala Glu Ile Thr Arg Ile	
35 40 45	
CCT CTG TAC AAA GGC AAG TCT CTG AGG AAG GCG CTG AAG GAG CAT GGG	192
Pro Leu Tyr Lys Gly Lys Ser Leu Arg Lys Ala Leu Lys Glu His Gly	
50 55 60	
CTT CTG GAG GAC TTC CTG CAG AAA CAG CAG TAT GGC ATC AGC AGC AAG	240
Leu Leu Glu Asp Phe Leu Gln Lys Gln Gln Tyr Gly Ile Ser Ser Lys	
65 70 75 80	
TAC TCC GGC TTC TCA GAC AAC CAG CGG CTC TTC AAT AAT GCA GTC ATT	288
Tyr Ser Gly Phe Ser Asp Asn Gln Arg Leu Phe Asn Asn Ala Val Ile	
85 90 95	
CGT GTA CAA CAC CTG CAC CAG CTG GCT GCA AAA ATG ATT AAC GAC TTT	336
Arg Val Gln His Leu His Gln Leu Ala Ala Lys Met Ile Asn Asp Phe	
100 105 110	
GAG GAC AGC CTG TTG CCT GAG GAA CGC AGA CAG CTG AGT AAA ATC TTC	384
Glu Asp Ser Leu Leu Pro Glu Glu Arg Arg Gln Leu Ser Lys Ile Phe	
115 120 125	
CCT CTG TCT TTC TGC AAT TCT GAC TAC ATT GAG GCG CCT GCT GGA AAA	432
Pro Leu Ser Phe Cys Asn Ser Asp Tyr Ile Glu Ala Pro Ala Gly Lys	
130 135 140	
GAT GAA ACA CAG AAG AGC TCT ATG CTG AAG CTT CTT CGC ATC TCT TTT	480
Asp Glu Thr Gln Lys Ser Ser Met Leu Lys Leu Leu Arg Ile Ser Phe	
145 150 155 160	
CAC CTC ATT GAG TCC TGG GAG TTC CCA AGC CAG TCC CTG AGC GGA ACC	528
His Leu Ile Glu Ser Trp Glu Phe Pro Ser Gln Ser Leu Ser Gly Thr	
165 170 175	
GTC TCA AAC AGC CTG ACC GTA GGG AAC CCC AAC CAG CTC ACT GAG AAG	576
Val Ser Asn Ser Leu Thr Val Gly Asn Pro Asn Gln Leu Thr Glu Lys	
180 185 190	
CTG GCC GAC TTG AAA ATG GGC ATC AGT GTG CTC ATC CAG GCA TGT CTC	624
Leu Ala Asp Leu Lys Met Gly Ile Ser Val Leu Ile Gln Ala Cys Leu	
195 200 205	
GAT GGT CAA CCA AAC ATG GAT GAT AAC GAC TCC TTG CCG CTG CCT TTT	672
Asp Gly Gln Pro Asn Met Asp Asp Asn Asp Ser Leu Pro Leu Pro Phe	
210 215 220	

GAG	GAC	TTC	TAC	TTG	ACC	ATG	GGG	GAG	AAC	AAC	CTC	AGA	GAG	AGC	TTT	720
Glu	Asp	Phe	Tyr	Leu	Thr	Met	Gly	Glu	Asn	Asn	Leu	Arg	Glu	Ser	Phe	
225					230					235					240	
CGT	CTG	CTG	GCT	TGC	TTC	AAG	AAG	GAC	ATG	CAC	AAA	GTC	GAG	ACC	TAC	768
Arg	Leu	Leu	Ala	Cys	Phe	Lys	Lys	Asp	Met	His	Lys	Val	Glu	Thr	Tyr	
			245					250						255		
TTG	AGG	GTT	GCA	AAT	TGC	AGG	AGA	TCC	CTG	GAT	TCC	AAC	TGC	ACC	CTG	816
Leu	Arg	Val	Ala	Asn	Cys	Arg	Arg	Ser	Leu	Asp	Ser	Asn	Cys	Thr	Leu	
			260					265					270			
TAG																819

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 272 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Arg	Gly	Ser	His	His	His	His	His	His	Gly	Met	Ala	Ser	Met	Thr	
1				5					10					15		
Gly	Gly	Gln	Gln	Met	Gly	Arg	Asp	Leu	Tyr	Asp	Asp	Asp	Asp	Lys	Asp	
		20						25					30			
Pro	Ser	Ser	Arg	Ser	Ala	Glu	Ile	Gly	Ser	Ala	Glu	Ile	Thr	Arg	Ile	
		35					40						45			
Pro	Leu	Tyr	Lys	Gly	Lys	Ser	Leu	Arg	Lys	Ala	Leu	Lys	Glu	His	Gly	
	50					55					60					
Leu	Leu	Glu	Asp	Phe	Leu	Gln	Lys	Gln	Gln	Tyr	Gly	Ile	Ser	Ser	Lys	
65					70					75					80	
Tyr	Ser	Gly	Phe	Ser	Asp	Asn	Gln	Arg	Leu	Phe	Asn	Asn	Ala	Val	Ile	
			85						90					95		
Arg	Val	Gln	His	Leu	His	Gln	Leu	Ala	Lys	Met	Ile	Asn	Asp	Phe		
		100						105				110				
Glu	Asp	Ser	Leu	Leu	Pro	Glu	Glu	Arg	Arg	Gln	Leu	Ser	Lys	Ile	Phe	
	115						120					125				
Pro	Leu	Ser	Phe	Cys	Asn	Ser	Asp	Tyr	Ile	Glu	Ala	Pro	Ala	Gly	Lys	
	130					135					140					
Asp	Glu	Thr	Gln	Lys	Ser	Ser	Met	Leu	Lys	Leu	Leu	Arg	Ile	Ser	Phe	
145				150						155					160	
His	Leu	Ile	Glu	Ser	Trp	Glu	Phe	Pro	Ser	Gln	Ser	Leu	Ser	Gly	Thr	
			165					170						175		
Val	Ser	Asn	Ser	Leu	Thr	Val	Gly	Asn	Pro	Asn	Gln	Leu	Thr	Glu	Lys	

180 185 190

Leu Ala Asp Leu Lys Met Gly Ile Ser Val Leu Ile Gln Ala Cys Leu
195 200 205

Asp Gly Gln Pro Asn Met Asp Asp Asn Asp Ser Leu Pro Leu Pro Phe
210 215 220

Glu Asp Phe Tyr Leu Thr Met Gly Glu Asn Asn Leu Arg Glu Ser Phe
225 230 235 240

Arg Leu Leu Ala Cys Phe Lys Lys Asp Met His Lys Val Glu Thr Tyr
245 250 255

Leu Arg Val Ala Asn Cys Arg Arg Ser Leu Asp Ser Asn Cys Thr Leu
260 265 270


